

ENGLISH AS A MEASUREMENT STANDARD IN THE NATIONAL EXAMINATION: SOME GRASSROOTS' VOICE

Gunadi Harry Sulistyo

Universitas Negeri Malang, Indonesia

Abstract: The National Examinations, including English as one component, have always implored controversies. And the government has still been determined in its commitment to execute the national examinations backed up with several accompanying arguments as an attempt to 'smarten the life of the nation'. This has not, however, likely been well understood by all parties with a view to gaining communal supports from all. A small survey was conducted to analyze English teachers' point of view, as frontline agents in the field, concerning issues related with the national examinations. It turns out that in spite of their positive viewpoint in the role of tests, when challenged to accepting responsibility in the context of classroom-based assessment as a part of their autonomous roles, they demonstrate an impression of disinclination.

Key words: national examinations, teachers, view

The government's effort in improving the quality of Indonesia's human resources through national education has been undertaken all the way through numerous routes (see for instance, *Kebijakan Direktorat Pendidikan Lanjutan Pertama*, 2004, *Panduan Program Peningkatan dan Pengembangan Pendidikan Daerah*, 2005, and *Rencana Strategis Direktorat Pendidikan Lanjutan Pertama*, 2005). One is through the management of national examinations. However, not all the government policies in managing national examinations have moved along smoothly as expected. The launch of the decree of the Minister of National Education No. 153/U/2003, 14 October 2003 concerning National Examinations (henceforth *UN – Ujian Nasional*) for the academic year 2003/2004 invoked an intense controversy. Public reactions radiating that time echoed negations. Husnawati (2004), for instance, criticized *UN* as having seri-

ous flaws. *UN* has been considered potential in leading students to mechanistic rote learning, with the sole purpose in response to limited learning targets while other potentials are neglected. Besides, the establishing of the passing grade (for the upcoming *UN*, it is set up at a figure 4.5) constitutes a stressor that potentially horrifies students. Also, it is argued, test items in *UN* are prone to invoke thinking that hinders the development of the students' multiple intelligences. Still in the same wavelength as Husnawati's, Santoso (2004) has been of the negative opinions as follows. Test items in *UN* have merely oriented to academic pedagogical perspectives, building a possible character for the students to be insensitive to their environment. Also, the test items in *UN* do not stimulate students' logical thinking, innovative thoughts, and emotional survival. Not just that public concern, a more academic discussion to address *UN* was also held. In a seminar held in State University of Jakarta on 11 April 2004, for example, negative reactions were also observed. It was among other things stated that *UN* has seized the students' rights. Even further, in a hearing between the representatives of the Department of National Education and *Komisi VI* that deals with education matters and affairs, negativity was also aware of. Even, there came up a strong push from the representatives to put *UN* aside from the evaluation system of students' learning in the upcoming academic years. The reactions have been, first, the *passing grade* has set the students at a point they cannot negotiate but must accept. Next, it is believed that the teachers' quality has not been adequately satisfactory as a result the students' readiness to sit in the examinations is still bare minimum. Also, viewed from the budget, it uses up a large financial resource. And most importantly, it is argued, *UN* contravenes *UU Sisdiknas no 20, 2003* (Jawa Pos, 5 May 2004).

Apart from these disapproving points of view, some more positive thinking on the implementation of *UN* is also observed. Musthofa (2004), for instance, believes that everyone needs to support the implementation of *UN* for several reasons that follow. In the first place, *UN* constitutes an idealized measure in the national vision that is expected to be able to get rid off the image regarding the low quality of national education in this global era. In addition, it is argued *UN* plays also an important role as a motivating force for the students to envision in their learning. It is also a means for the students to cultivate a moral fibre of fair competition. Not just the students, the teachers will also gain benefits from the implementation of *UN*. *UN* will stimulate and motivate teachers to conduct a better instruction in the classroom. Further, the decision to include three subjects to be administered in *UN*: English, Indonesian,

and Maths, is considered a strategic policy in responding to the challenges in the global era. Similar to Musthofa's viewpoints, Prameswari (2004) has also been of the opinion that *UN* will constructively bridge the gap existing among schools and that among different districts or provinces if all educational stakeholders have a strong commitment to implement *UN*. Besides, *UN* is expected to play an initial momentum that can be used as a springboard to improve the quality of national education practices.

National examination like *UN* has been implemented formally since the academic year 2002/2003. The government policy in its firm commitment to implement *UN* is in no way educationally unsubstantiated. As has been reiterated by the then Director General of Primary and Secondary Education (2004:2), *UN* comprises an integral part of a system implemented to improve the quality of national education that has still been considered representing a dark side. It is further assured that in a curriculum system the students' learning based on the curriculum calls for evaluation. Examinations represent attempts to measure the students' achievement on the pre stated goals outlined in the curriculum. This is inevitably a form of accountability measures. In addition, evaluation is a part of a system that is useful for the enhancement of teachers' teaching and students' learning processes. In this way evaluation has corrective measures that will have backwash effects on the performance of the teachers and students alike.

The rationale launched by the Director sounds to be educationally normative. Nevertheless, the reasoning is substantial in that it is actually a form of elaboration of education concepts put into actual practice. Theoretically, several interrelated aspects are dealt with in the realm of instructional strategies one of which is testing (Dick and Carey, 1996:191). Of the strategies, the over-exposed aspect that is publicly known is instruction. The others, aware or not, are commonly not well recognized and are considered 'an attachment', minority in importance. However, actually evaluation plays a substantial role in education although many have still understood it partially (Athanasou and Lamprianou, 2002:1).

As a state institution responsible for the quality of national education, the Department of National Education with its all constituents have the responsibility to always make every attempt to 'smarten the life of the nation' as has been mandated in the Constitution 1945 - *Undang-Undang Dasar 1945*. This follows then that examinations as an integral part of evaluation system as argued

previously can actually be well understood by all parties with a view to gaining communal supports from all.

The discussion aforementioned implies apprehensions in the outlook and expectation between the Department of National Education as the representative of the government that is mandated to manage the development of national education and the citizens – the ones who are actually the target audience of the development. Such a tension may persist on and on, thus potentially resulting in unfruitfulness in terms of necessary actions to contend with the situation. To provide a fuller context of the situation, it seems desirable to capture empirical broader views other than those coming from the bureaucrats and the public. Other voices of the practitioners in the classroom – teachers - need also to be heard.

PROCEDURES

The purpose of the writing is to present an empirical sketch in the form descriptive report on a preliminary investigation into what views and opinions teachers as a front liner in the classroom have in conjunction with the implementation of *UN*. The conduct is preliminary in that it only includes a limited number of English teachers of M.Ts in several provinces: East Java, Central Java, Banten, South Kalimantan, East Kalimantan, West Sumatera, and Bangka Belitung. In total, there were 24 English M.Ts teachers involved (13 males and 11 females). They were invited as participants of National Certification Program organized by English Department of Faculty of Humanities and Culture, Islamic State University of Malang in 2005.

In the program, the participants were certified for the teaching of English at M.Ts. because the major they had had prior to the certification program was of non English education background while in their schools they were formally assigned to teach English. To bridge the gap existing in the mismatch between the major they have and the official assignment in their school, the certification program was run.

Most of the subjects (79%) held an S-1 degree in which several majors were recognized: guidance and counseling, English literature, Islamic education, education of business management, chemistry education and others. Four (16.7%) held a Diploma 3 certificate, and one (4.2%) had a Diploma 2 certificate. Their teaching experiences range from 0.5 year to 38 years with mostly (62%) having 5 years teaching experiences or below, four (16.7%) between 5.1

to 10 years, and five participants (20.8%) having 10 years teaching experiences or above.

To collect data related with the implementation of *UN*, a set of questionnaire was established. The questionnaire encompasses their views and opinions among other things that deal with the role of *UN* to improve the students' competencies in English, a choice to depend on *UN* or teacher-made evaluation, academic advantages, the English skills and components to be examined, the function of scores derived from *UN*, and their readiness. The questionnaire takes the form of question with a dichotomous response (Yes – No), extended with a probing open-ended format for further explanation on the choice (Yes – No) considered (Nitko, 1996). Peer validation was made on the questionnaire, resulting on the changes of the formulation of statements deemed to be potentially confusing to respondents.

The data which were the participants' opinions and views were collected at the first meeting of one of the courses, 'Language Test Construction: Theory and Practice'. They were allotted 15 minutes to respond to all questions posed in the questionnaire, exactly the last 15 minutes before the first meeting ended. To reveal the participants' opinions and views, quantitative data were then analyzed descriptively to find the frequency of occurrences and its corresponding percentage. For the purpose, *SPSS version 10.0 for Windows* was applied to analyze the quantitative data. Data other than these were analyzed qualitatively to reveal the patterns of relation among verbal responses made by the respondents.

RESULTS

The results of the analysis addressing the data concerning the participants' belief whether the inclusion of English as a component of national examinations leads to the promotion of the students' competencies in English are presented in Table 1.

The data in Table 1 indicate not all participants come to a single agreement approving that the inclusion of English as a component of national examinations will automatically go along with the improvement of the students' learning of English. Nevertheless, the majority (75%) holds the views while the minority (25%) has the opposite beliefs.

Table 1. English in National Examinations to Improve Students' Competencies in English

Response	Frequency	Percent	Valid Percent	Cumulative Percent
No	6	25.0	25.0	25.0
Yes	18	75.0	75.0	100.0
Total	24	100.0	100.0	

The reasons put forward by those who favor the belief in the improvement of students' learning, however, vary. Mostly relate with the students' learning. They believe that the policy of including English in *UN* will motivate students to learn English better, one form of which will be that the students will join English courses outside the formal class sessions. In a close tendency to this view, some believe that *UN* sets a student's profile against national standards as a follow up treatment after classroom evaluation.

The second reason favored by one-thirds of the participants refers to the role of English as an international language. They view that English is an important international language needed much in the global era. To them, English functions as a life skill to master in the global era. It is an important tool for acquiring science and technology. In their vision, *UN* is the tool that is expected to function as a whip that can enhance the English learning to the global standards.

The third reason relates with the teachers. To them, *UN* is a national standard for the teachers to aspire. Close to this teacher-oriented outlook is the view envisioned by a participant stating that *UN* will function as a valid source of information on which to base evaluation on the teachers' professionalism, i.e. their quality in teaching performance. Respondents in this group are of the opinion that teachers will be motivated to teach English better to the level required by the national standards. This view echoes the quality teaching performance of the teachers. In terms of quantity, one is of the opinion that teachers will provide additional teaching hours.

Other opinions launched by single respondents deal with the technicalities in the scoring used in *UN* and the involvement of parents. To them, *UN* is more objective in scoring than the teachers in general; whereas with *UN* parents will be challenged to assist their children in the learning of English.

Those aforementioned are the views from the respondents who are in favour of *UN*. Although not many respondents belonging to the minority group, it seems that their outlook is worth documenting. The aspects in the views that support *UN* include a wider spectrum such as students, teachers and others. However, in the minority group the aspects seem to be limited. Most relate with the testing practices that *UN* holds. For example, one criticizes that English as a means of communication is not properly assessed. This view is congruent with another one thus, providing a logical consequence, stating that the multiple-choice format used in *UN* leads students to speculation in responding to test items. To this group, then, the test items do not measure accurately for there are good learners but do not pass English in *UN*. Due to the test format also, it is viewed that *UN* does not guarantee skilful communication in English. Further more, in *UN* not all language skills are examined. As a result, those who pass English in *UN* are not yet able to ‘produce’ English as a means of communication.

Another aspect touched on in the minority group is concerned with teachers’ teaching practices. They believe that *UN* makes teachers tend to teach English materials as a target, but not necessarily under the spirit of mastery learning – a need to learn English as a means of communication. Related with this view is that, as a result, students will tend to learn for marks rather than functional skills of English to communicate. Under a resourceful teacher, it is believed that a test is not just a single means that can be used to determine one’s English ability.

The second item in the questionnaire asks the teachers’ preference in evaluating students’ final learning achievement: to depend on the tests developed by the teachers themselves or the government. The results of the data analysis in this aspect are presented in Table 2.

Table 2. Teachers’ Preference: Dependence or Independence in Evaluating

Responses	Frequency	Percent	Valid Percent	Cumulative Percent
Government	15	62.5	62.5	62.5
Self	9	37.5	37.5	100.0
Total	24	100.0	100.0	

Table 2 above clearly shows that the majority (62.5%) prefers to depend on the tests provided by the government; only a small number (37.5%) want to evaluate the students' achievement using their own tests (teacher-made tests). The teachers' dependence on the government is based on the following conditions as they admitted. In the first place, the test developed by the government is envisaged to have better objectivity although they cast a doubt concerning the content validity of the test developed by the government.

The reason that the test developed by the government has better objectivity is possibly an explanation for another genuine response admitting the teacher's inability (*yet?*) to construct good tests and the avoidance of unfairness possibly committed by schools in determining those passing *UN* just for the purpose of 'school prestige' by admitting a high percentage of the school graduates. On the other side, the teacher's doubt on the content validity seems to be understandable for, as has been revealed in their view, not all language skills might be proportionately assessed in *UN*.

Next, a more idealized response states that they prefer to depend on the government-made test because, they believe, the test reflects national standards. As such, they argue, the implementation is directed to two conditions: a national standardization of examining English learning outputs the learning processes of which have been mandated to schools, and a control to evaluate the teachers' teaching which is necessarily a form of quality assurance on teachers' performance. Thus, the latter is something like a system to portray the teachers' professional competence in teaching. Another piece of idealized argument states that *UN* is believed to establish a normative motivating vigour for English teachers to head to become professional based on national standards.

Those who believe that the teacher-made assessment has a more positive side hold their view as follows. The majority claims that teachers know students' better in terms of their students' ability and learning progress so that in assessment the test design matches the teaching. This claim reiterates other assertions maintaining teachers' content validity: the relevance of what has been taught and what to test, and perceived unfairness of *UN* for only including a limited scope of materials. What worries them further is the state that students' potentials should not be solely examined from the results of *UN* at the expense of other potentials not revealed in the *UN*.

Other points of arguments made by those in favour of teacher-made tests encompass the importance of the learning process. They state that language

ability cannot be just judged on the basis of single testing; learning process-based evaluation is more important and valid. Further, with a view to mastery learning, they also assert that teaching and learning processes are more important not just the final examination.

Another area of concern addresses the nature of the test format and a single administration in *UN*. Due to the limited response format used in *UN*, i.e. multiple-choice one (Purpura, 2004; Nitko, 1996), there is a perception of greater likelihood of guessing on the part of the students in completing the test. Also, evaluating students' learning progress just on a single examination can be unfair enough and not parallel: first, with a view that language is performance not just static knowledge and second, with regards to existing conditional states of the student, say for example illness or other mental or physical limitations. All these follow that they suggest teachers should be granted freedom to evaluate their students whenever required, let alone on the consideration that each school's performance varies considerably from one area to another.

Finally, teacher-made test orientations sustain a professional growth claim on the part of the teachers. They assert that teacher-made tests stimulate teachers' creativity in constructing tests and evaluating the students' learning output. In this view, teachers need to put their expertise, they claim, in test construction and evaluation. They are not just instruction deliverers. Next, they argue that teachers can give maximum standards of mastery not just the minimal like *UN*.

Apart from those conflicting views described above, there is one interesting vision. This idea emphasizes the importance of pronouncing a win-win solution in which teachers and the government can share the responsibility in the process of assessing and evaluating the students learning progress and learning output. This view is in line with an assertion that educational stakeholders comprise necessarily both central and local educational agents, including teachers.

The third item in the questionnaire elicits the teacher respondents' opinions concerning the benefits they may gain as an English teacher with English as a component of *UN*. The results of the data analysis on the teachers' responses are presented in Table 3.

The table above demonstrates that those who feel they gain something from *UN* and those who do not are almost balanced (54.2% and 45% respectively). This piece of information seems to contradict a bit the data concerning the teachers' preference on using teacher-made tests over the ones from the government, which outweighs the second group (62.5%).

Table 3. Prospective Advantages Gained

Response	Frequency	Percent	Valid Percent	Cumulative Percent
No	11	45.8	45.8	45.8
Yes	13	54.2	54.2	100.0
Total	24	100.0	100.0	

Indifference is observed in their response. Three state that there is no academic advantage for them as they are private school teachers. This response, while indicating lack of academic concern, is in line with another observation affirming that with *UN* teachers will be administratively busier than when there is no *UN* – which in their view the time taken up will be more useful and better for making teaching preparations.

A more positively academic outlook is also observed. This is related with the teachers' teaching practices in the classroom. First, teachers will passively depend on the results of *UN*, not other aspects of evaluation. What further makes them troubled concerns validity, in that as *UN* technically covers a limited scope of English aspects to be tested (normally reading and writing), students will focus only on what is tested with the emphasis on correctness; other English functional skills neglected. This situation is potentially to be worsened by the classroom practices where the teachers' teaching activity in the classroom is geared to analyzing test items like those in *UN*, rather than stimulating students' interest, guiding and facilitating students' learning. For some confident schools, *UN* standards of passing the level set up are viewed to be soul-destroying. *UN* merely applies minimum performance; as a result, some good students will tend to learn slightly rather than make every attempt for demonstrating their best abilities.

On the other hand, those who feel they gain something from *UN* launch their arguments as follows. First, as has been also revealed previously, the advantage deals with teachers' professionalism. Most (6) believe that with *UN* teachers are posed with a challenge to upgrade their professionalism by providing the best services to students. The results of *UN* will constitute useful feedback to the teachers. As a consequence, teachers will know their weaknesses and strengths in their teaching towards standards set up in *UN*. When this is achieved, then the function of *UN* is fulfilled accordingly in that *UN* is a challenge to improve quality in education.

Next, to some *UN* has a function of utility. Teachers use the image of *UN* as a tool to motivate students to learn English better. Compared to the condition of non existence of *UN*, with *UN* it is easier for the teacher to motivate students to learn English as English is one subject in *UN*. To school masters, however, *UN* gives an advantage in another way. Teachers feel that their headmaster tends to give more attention to English as a subject in *UN*, including the English teacher. For example, the provision of English textbooks becomes the school priority.

Apart from these arguments, personal and pragmatic advantages are also observed. For example, *UN* is the opportunity for the teacher to offer and provide additional lessons meaning that more teaching hours will be more income. Another advantage is that as the tests are supplied, schools have no need to prepare test items.

Content validity in *UN* in terms of language skills and components to be tested has become an issue. This has also been true with the results of the present study as shown in the previous data presentation. The next question in the questionnaire tries to probe the teachers' judgement on language skills and components needing to be considered in *UN*. The results of the analysis on the data collected are presented in Table 4.

Table 4. *UN* Components to Be Assessed

Response	Frequency	Percent	Valid Percent	Cumulative Percent
All	7	29.2	29.2	29.2
Particular	17	70.8	70.8	100.0
Total	24	100.0	100.0	

Disagreement is observed in Table 4 as to the teachers' response to the question whether it is necessary to include all or particular aspects of English in *UN*. Quite contradictory to the issue of content validity, most teachers (70.8%) consider it necessary to include only particular aspects of English in *UN*; the rest (29.2%) all aspects.

The particularity of the aspects of English in *UN* also demonstrates variability. Most, however, are of the opinion of the necessity to focus on integratedness of the language skills (listening, speaking, reading and writing: L-S-R-W following the order of teaching tradition of the Audio Lingual Method) to be

tested in *UN* with a differing emphasis. Some believe that reading and writing need to be prioritized; others speaking and listening. The general argument they put forward is that English as communication necessitates language skills used functionally. In their words, language skills ‘challenge the students to learn to communicate functionally and the teacher can assess their learning’; ‘these [language skills] challenge the students to be active practising their English’ and ‘students need more practice not just know a lot about language.’

With a view to language skills still but without vision of integrativeness, individual respondents perceive that these language skills: listening, speaking, reading, and writing need to constitute *UN* materials. The order they propose varies however: (1) speaking, listening, writing, and reading (S-L-W-R) with a focus on speaking skills because, it is argued, speaking is the benchmark that someone has the mastery of English; (2) reading, writing, listening (R-W-L), and (3) speaking, listening, writing, reading (S-L-W-R) as what normally happens in *UN*.

Others consider technical practicality in the *UN* administration. Thus, the order reading, writing, grammar, and vocabulary (R-W-G-V) or reading, writing, vocabulary, and grammar (R-W-V-G) with an emphasis on grammar and reading is easier to be administered. Another one considers it easier: reading, speaking, writing, listening (R-S-W-L). Also, the composite reading, vocabulary, and grammar (R-V-G) is easier to be administered than writing, listening, and speaking in which their scoring tends to be subjective.

Another view, which tends to be language component-oriented, sees the importance of priority. In this view, vocabulary, reading and grammar (V-R-G) are testable in *UN* while writing, speaking, listening (W-S-L) are assessed in teacher-made tests.

Unlike the ‘particular’ view of *UN* components, the ‘all’ view of *UN* components can further be sub grouped into two: one with the focus on vocabulary, and the other one on speaking. Surprisingly, however, only one sees the necessity to test all components in an integrative fashion in *UN*.

Opposite to the previous question, the next question asked deals with the kinds of English aspects to be assessed when the decision to let the students pass *UN* is granted autonomously to the classroom teachers. In other words, the question forwarded to the teachers is if the teachers are given the authority to evaluate the students’ learning achievement and certify their English competency, what English aspects are essentially to be assessed? The results of the data analysis on this variable are presented in Table 5.

Table 5. Language Aspects to be Assessed by Teachers

Response	Frequency	Percent	Valid Percent	Cumulative Percent
All	18	75.0	75.0	75.0
Particular	6	25.0	25.0	100.0
Total	24	100.0	100.0	

As shown in the table, the majority (75%) of the teachers are in favour of comprehensive assessment: measuring all English aspects in the teacher-based assessment. This means that all language skills (L-S-R-W) and language components (Grammar and Vocabulary), they view, need to be evaluated to certify the students in their English competencies upon completing their study. This response is clearly contradictory to the previous findings on the *UN*-based assessment in which the majority of the teachers as the respondents (70.8%) demonstrate their preference to include only particular English aspects in *UN*. These two data sets actually imply congruence and consistency of the respondents in that they seem to have the tendency to include all English aspects in the teacher-based assessment and on the contrary particular English aspects in *UN*.

When they were asked further, why and what to prioritize, they (50%) state that ability in English means functional competencies to communicate in all English skills: listening, speaking, reading and writing integrated with grammar and vocabulary. In terms of priority, most put a more emphasis on the solid mastery of English speaking skills. They set eyes on the speaking ability as a benchmark representing the ability in English, speaking with confidence albeit inaccurate. With regards to the order of importance, their opinion can be classified into two major almost identical streams: both (a) the stream S-L-R-W-G-V and (b) the stream S-R-L-W-G-V. Other proposed choices include R-V-G-L-S-W, R-W-S-V-G-L, and V-R-W-G-S-L.

Beyond these responses, an interesting view is observed in relation to the management of the assessment of all these English aspects. Considering the number of the English aspects to be assessed, one respondent offers a suggestion to assess these aspects; that is, the teachers may manage the time of assessment so that the assessment can be spread over a certain period, not necessarily a simultaneous administration.

The responses from those in favour of assessing particular aspects of English mainly reveal the need to limit to the testing of the four English skills (L-S-R-W) with language components integrated into these skills. Two main streams are also observed: (a) S-R-W-L or S-L-R-W and (b) R-W-L/S. The rationale they propose for testing these skills is two folds: English as communication and the availability of limited time

As has been revealed in Table 2 previously, the majority of the respondents (62.5%) tend to rely on the government tests. However, when the teachers are probed further concerning their standing if the government leaves the evaluation to them alone with assurance that the results of evaluating students really reflect genuine competencies, all (100%) responded positively. They state that the mandate awarded to them autonomously to assess the students is perceived as a challenge to deal with.

The justification as regards with the mandate, it is revealed, varies considerably. As presented in Table 6, the results of the analysis specify the data.

Table 6. Rationale for the Teachers' Standing

Response	Frequency	Percent	Valid Percent	Cumulative Percent
Personal moral responsibility	5	20.8	20.8	20.8
Professional responsibility	3	12.5	12.5	33.3
School prestige & personal moral responsibility	1	4.2	4.2	37.5
Personal prestige & professional responsibility	1	4.2	4.2	41.7
Personal moral & professional responsibilities	4	16.7	16.7	58.3
School prestige, personal moral & professional responsibilities	1	4.2	4.2	62.5
Personal prestige, personal moral & professional responsibilities	1	4.2	4.2	66.7
School prestige, personal prestige, personal moral & professional responsibilities	8	33.3	33.3	100.0
Total	24	100.0	100.0	

As demonstrated in Table 6, the acceptance of the challenge can be seen from the number of sources of drive. When observed from the single drive, the majority of the teachers (20.8%) accept the responsibility because of personal moral responsibilities, which is then followed by professional responsibilities (12.5%) However, when seen from the combination of the drive, the majority (33.3%) accept the challenge because of four sources of drive in combination:

school prestige, personal prestige, personal moral & professional responsibilities. This is then followed by the combined drive: personal moral & professional responsibilities. So, it seems that personal moral & professional responsibilities collectively underlie the teachers' drive in accommodating the challenge from the government when it is really granted to them.

Acceptance of responsibility to assess the students' learning output on the teachers' own implies every effort made by the teacher with a view that the evaluation on the students' learning the teachers make really reflects genuine students' competencies. When they were asked what they would do to accept the responsibility, the teachers' responses turn out to vary considerably. Some reflect professional efforts while most (62.5%) are of personal interest. Of the professional undertaking, most claim that they will vary their teaching methods (4), which are then followed by an endeavor to have the students have more practice on language skills (3). In addition, some maintain that they will use varied means of instructional media (2) in their teaching learning process. Some others feel the need to make the time allocation optimally met (2) already set up by the government. Apart from these determinations, other individual teachers consider it necessary to perform several things as follows: upgrading themselves by joining short courses, motivating the students to do their best, running a school program like an English day, and surprisingly by collaborating with other course services providers.

The other efforts claimed by the teachers to be made are mostly related with providing additional assistance beyond the classroom practices. These include running a private course and providing more hours for enrichment of students' instructional learning (15).

The next item in the questionnaire solicits the teachers' responses on their position concerning their endorsement to the use of UN scores for one of the criteria to pass the students. The results of the data analysis on this matter are presented in Table 7.

Table 7. UN Scores as One Criterion to Pass Students

Response	Frequency	Percent	Valid Percent	Cumulative Percent
No	17	70.8	70.8	70.8
Yes	7	29.2	29.2	100.0
Total	24	100.0	100.0	

The data in Table 7 inform that the majority of the teachers (70.8%) come to the disagreement for using the scores obtained from *UN* as a criterion to pass the students upon their completion of their study. The teacher's response is consistent when they were further invited to pose their view concerning the use of *UN* scores for non pass purposes. The data on this variable is presented in Table 8.

Table 8. *UN* Scores for Non Pass Purposes

Response	Frequency	Percent	Valid Percent	Cumulative Percent
No	2	8.3	8.3	8.3
Yes	22	91.7	91.7	100.0
Total	24	100.0	100.0	

As shown in Table 8, even more teachers (91.7%) view that the students' scores obtained from *UN* are usable for non pass purposes and only a small number (8.3%) are of the opinion that the *UN* scores are for pass purposes. Their view, however, diverge to a large extent when they were solicited further concerning the use of the students' scores obtainable from *UN* for non pass purposes. The data on this matter is presented in Table 9.

Table 9. Use of *UN* Scores for Non Pass Criteria

Response	Frequency	Percent	Valid Percent	Cumulative Percent
Absence	2	8.3	8.3	8.3
A	2	8.3	8.3	16.7
B	10	41.7	41.7	58.3
C	5	20.8	20.8	79.2
A and B	2	8.3	8.3	87.5
B and C	1	4.2	4.2	91.7
A, B, and C	2	8.3	8.3	100.0
Total	24	100.0	100.0	

Note:

Absence : no response is evidenced from subjects

A : Basis for improvement of school effectiveness performance

B : Basis for improvement of teachers' professional performance

C : Others: remediation.

Table 9 clearly presents the fact that almost half of the respondents (41.7%) view the importance of improvement of teachers' professional performance based on the students' score obtained from UN. Some others consider it necessary for using students' *UN* scores for the purpose other than improvement of school effectiveness performance and improvement of teachers' professional performance, for instance, remedial teaching. Surprisingly, only a small number (8.3%) of the teachers think that the students' *UN* scores can be useful as a basis for improvement of school effectiveness performance. Similarly, only a small number (8.3%) of the teachers are of the opinion that the students' *UN* scores can be useful as a basis for both improvement of school effectiveness performance and improvement of teachers' professional performance as well as the combined use of the scores for the basis for both improvement of school effectiveness performance and improvement of teachers' professional performance and other purposes.

UN is, as its name suggests, to be applied into operation nationally. This means that all schools, private and public, in all districts of differing circumstances, are to employ *UN* as the government policy without exception. This potentially invites public issues of unfairness. To address such an issue more objectively, a question designed to collect the data on fairness is constructed. Therefore, the questionnaire also forwards a question probing the teachers' view concerning the fairness of *UN* as a test. The results of the analysis on the data are presented in Table 10.

Table 10. Unfairness of *UN* as a Test in Practice

Response	Frequency	Percent	Valid Percent	Cumulative Percent
No	5	20.8	21.8	21.7
Yes	18	79.2	79.2	100.0
Total	24	100.0	100.0	

As demonstrated in Table 10, the majority of the teachers as respondents (79.2%) are of the view that *UN* as a test is not fair practically. The rationale they put forward for their view differ substantially. However, most (50%) are of the opinion that an examination conducted nationally, *UN* is no good judgment in the context of teaching. For instance, schools in different areas obviously differ qualitatively and quantitatively in terms of their teaching-learning

facilities, teachers as human resources, teaching-learning processes, and students as input. The condition is worsened by the confusion over the implementation of the curriculum envisioned to be stuck to. Also, they believe that the central stakeholders are less sensitive to the teaching-learning conditions in districts in which the national curriculum is not yet well implemented. Aside from these, other views relate with testing practices. Some state that *UN* covers only a limited scope, mostly reading comprehension while speaking is not covered. This among other things follows that many important aspects and potentials attributed to by the students are not addressed accurately in *UN*. Another view is that the multiple-choice format used in *UN* encourages poor-level students to perform blind-guessing. Also, *UN* is a one-shot evaluation the pass standard of which is potential to contemporary threats to students when they are sitting on the examination such as distress, test-illness, and the like.

UN-based evaluation may be contrasted with classroom-based evaluation, which to some degree implies teacher-based evaluation (cf. Stiggins, 1994; Section 58, Verse 2, Chapter 16, *Undang-Undang RI No. 20, 2003*). When teachers' preference tends to go to the classroom-based evaluation, teachers certainly assume more responsibility accordingly. Three questions that are meant to address this issue were established to look into the teachers' readiness in assuming more responsibility, constructing quality items, and interpreting test scores accurately.

The results of the data analysis on the teachers' accountability are presented in Table 10, those on the teachers' readiness on constructing quality items in Table 11, and those on the teachers' readiness on interpreting scores accurately are presented in Table 12 correspondingly.

Table 10. Teachers' Claim on Accountability with Classroom-Based Evaluation

Response	Frequency	Percent	Valid Percent	Cumulative Percent
No	12	50.0	50.0	50.0
Yes	12	50.0	50.0	100.0
Total	24	100.0	100.0	

Table 10 evidently exhibit the fact that the number of those who are ready to assume responsibility with classroom-based evaluation is the same as

that of those who are not ready yet to do so, figuring each at 50%. This means that a half of the respondents are ready and the other half are not.

Table 11. Teachers' Claim on Readiness with Classroom-Based Evaluation on Constructing Quality Items

Response	Frequency	Percent	Valid Percent	Cumulative Percent
Not Ready Yet	19	79.2	79.2	79.2
Ready	5	20.8	20.8	100.0
Total	24	100.0	100.0	

With regards to their readiness on constructing quality items, the majority of the respondents (79.2%) as shown in Table 11 obviously maintain that they are not ready to assume the responsibility. Only a small number (20.8%) assert that they are ready. A similar claim is also observed on their readiness with classroom-based evaluation on score interpretation.

Table 12. Teachers' Claim on Readiness with Classroom-Based Evaluation on Score Interpretation

Response	Frequency	Percent	Valid Percent	Cumulative Percent
Not Yet Ready	18	75.0	75.0	75.0
Ready	6	25.0	25.0	100.0
Total	24	100.0	100.0	

Table 12 displays the fact that the majority of the respondents (75%) are not yet ready to take the responsibility for accurately interpreting students' scores derived from the test they make. Only a small number (25%) assert that they are ready.

In a nut shell, the results of the data analysis can be summarized as follows. The teacher respondents in the present survey are of the belief that the inclusion of English as a component of *UN* will be followed by the improvement of the students' learning of English. The reasons for this relate with the potential force of *UN* to motivate students' learning English. Next, the majority prefer to depend on the tests provided by the government. Besides, more teachers feel that they can benefit with English as a component of *UN*. Concerning con-

tent of *UN* in terms of language skills and components to be tested, most teachers consider it necessary to include only particular aspects of English in *UN*. However, on the teacher-based assessment, the majority are in favor of comprehensive assessment: measuring all English aspects. The acceptance of the mandate to have the teacher-based assessment, they view, is due to personal and moral responsibilities. The attempts they would take to accept the responsibility some reflect professional efforts while most are of personal interest.

The majority of the teachers disagreed at the idea for using the scores obtained from *UN* as a criterion to pass the students upon their completion of their study; they tend to agree to utilize scores obtained from *UN* for non pass purposes. Also, the majority are of the view that *UN* as a test is not fair practically. The teachers did not show a clear trend as to assuming responsibility with classroom-based evaluation. Finally, the teachers confess that they are not ready yet on both constructing quality items and accuracy in interpreting students' scores derived from the test they make when they are granted autonomy in evaluating their students' learning.

DISCUSSION

In terms of legal and normative measures, teachers are 'professionals whose responsibilities include planning learning processes, implementing them, and evaluating the learning outcomes, ...' (Verse 2, Section 39, Chapter XI, *Undang-Undang RI No. 20, 2003*; also Verse 1, Section 2, Chapter II, *Undang-Undang RI No. 14, 2005*). Pedagogically, teachers have been recognized to play an important role as a learning conductor, facilitator, motivator, and more recently as a reflective decision maker (Kellough and Kellough, 1999). In real life, however, particularly in national educational context, so far there has been a naïve and overwhelming perception that teachers are the front-liners, in that they are commonly considered the operational agents of change in the classroom. Teachers are expected to be classroom managers at work; but at the same time they are also ordinarily regarded as the implementers of the government policy in education. They seem to be didactically powerless in performing their roles and responsibilities in the classroom upon accommodating the policies coming from their super ordinates. Yet, the undeniable fact demonstrates that almost every day, teachers stand at the very front dealing with their students.

Teachers' viewpoints on *UN* as revealed in the present survey reflect that their autonomy in the classroom as secured by law and/or by profession has not been fully understood well by the majority. This may be an explanation why the teachers at schools as their work place seem to be at the crossroads when they are posed with policies felt to be compelling although these policies seem to contradict their beliefs on their educational practices. For example, the respondents believe that the inclusion of English as a component of *UN* will be followed by the improvement of the students' learning of English due to the potential force of *UN* to motivate students' learning English. While this view accords theoretical plausibility, in that tests have backwash effect both on students' learning behavior and teachers' conduct in instruction (Bachman, 1990), their other outlook denies this. They tend to believe the benefits of utilizing scores obtained from *UN* for non pass purposes.

From academic perspectives, it seems that teachers are of the view that improvement of students' learning quality can be enhanced through a test as a measuring device - though but not necessarily via *UN*, which, to them, seems to imply 'a more official load'. They on the face of it have this awareness of the importance of testing their language abilities. That is why two important points of view are theirs: the teacher-based assessment necessitates measurement of all English aspects and the teacher-based assessment is a symbol of personal and moral responsibilities. However, a genuine confession needing appreciation is that they are not fully ready to conduct classroom-based assessment in which automatically they will assume full responsibility. In so far as they still need expertise in language testing construction matters, they will, it is argued, leave the making of the tests to the government business.

The question that may be posed further is that are the phenomena as the findings of the present survey described above the characteristics exclusively limited to the sample of the present survey and not generalizable to a wider population? Affirmative answers may be understandable for the majority, not to mention all the respondents, have an inadequate encounter with matters related with language testing business. They are graduates of non English education institutions as their major. However, with a view to a hypothetical wider population, the phenomena described above may reflect a transition from centralized English learning testing practices to decentralized ones: a state of uncertainty. Many wish to conduct classroom-based assessment in the lack of solid and accurate knowledge and skills in language testing matters, particularly with those under the label of authentic assessment movement (O'Malley

and Pierce, 1996) as well as legitimacy. If this is the real case, then it is a challenge for all stakeholders in education to take several strategic measures if the business of testing students' English learning is considered to be of importance in equipping the students with the functional skills of English. Two modes are offered: improving their testing competencies as required in testing business, for instance through in-and-on service training on up to date assessment practices and launching a policy clarifying the role of UN and teachers' classroom-based assessment in the system of national evaluation.

CONCLUSION

The present survey has revealed a seemingly mixed state of voices from English teachers with non English education background on English as a sub-test in the national examinations – *UN*. Probably this state potentially reflects general pictures in the testing of English learning at schools. While there is recognition on the importance of testing and testing practices through the national examinations, there is a strong wish to conduct classroom-based assessment, which necessarily implies more demands on teachers' roles, responsibilities, and accountability on assessing their students' learning progress and achievement. In addition, however, there seems also to be a sense of insecurity in conducting classroom-based assessment to conclude the students' learning progress and achievement. This lack of confidence likely stems from insufficient knowledge and competencies in assessing students' English learning in the classroom using appropriate assessment methods and techniques, particularly the ones state-of-the-art. This articulation ought to be the concern of all stakeholders in education.

In response to such teachers' crucial need at schools, there seems to be a call for immediate endeavours on the part of all stakeholders in education, the government as well as institutions responsible for educating English teachers. Empowerment of teachers with functional and hands-on assessment know-how and consistent regulations to elucidate national testing practices at different levels of education units may be taken as the initiative in this matter. With this, in due course autonomy of teachers as a classroom evaluator will be in effect.

REFERENCES

- Athanasou, J. & I. Lamprianou. 2002. *A Teacher's Guide to Assessment*. Tuggerah, NSW: Social Science Press.
- Bachman, L.F. 1990. *Fundamental Considerations in Language Testing*. New York: Longman.
- Departemen Pendidikan Nasional. 2005. *Panduan Program Peningkatan dan Pengembangan Pendidikan Daerah*. Jakarta: Departemen Pendidikan Nasional.
- Departemen Pendidikan Nasional. 2005. *Ujian Nasional 2005/2006*. Jakarta: Departemen Pendidikan Nasional.
- Dick, W. & L. Carey. 1996. *The Systematic Design of Instruction*. New York: Harper Collins College Publishers.
- Direktorat Pendidikan Lanjutan Pertama. 2004. *Kebijakan Direktorat Pendidikan Lanjutan Pertama*. 2004. Jakarta: Departemen Pendidikan Nasional.
- Direktorat Pendidikan Lanjutan Pertama. 2005. *Rencana Strategis Direktorat Pendidikan Lanjutan Pertama*. Jakarta: Departemen Pendidikan Nasional.
- Husnawati, R. 2004. 'Peningkatan Semu Mutu Pendidikan.' *Jawa Pos*, 6 Mei 2004.
- Kellough, R.D. & Kellough, N.G. 1999. *Middle School Teaching: A Guide to Methods and Resources 3rd Edition*. Upper Saddle River, N.J.: Merrill, an Imprint of Prentice Hall.
- Musthofa, Z.A. 2004. 'Berfikir Global Bertindak Lokal.' *Jawa Pos*, 5 Mei 2004.
- Nitko, A.J. 1996. *Educational Assessment of Students. 2nd Edition*. Englewood Cliffs, NJ.: Merrill an Imprint of Prentice-Hall.
- O'Malley, J.M. & Pierce, L.V. 1996. *Authentic Assessment for English Language Learners: Practical Approaches for Teachers*. Reading, Mass.: Addison-Wesley Publishing Company.

- Purpura, J.E. 2004. *Assessing Grammar*. Cambridge, UK: Cambridge University Press.
- Prameswari, U. 2004. 'Langkah Awal Majukan Pendidikan.' *Jawa Pos*, 4 Mei 2004.
- Santoso, S.H. 2004. 'UAN itu Perlu, tapi ...'. *Jawa Pos*, 8 Mei 2004.
- Stiggins, R.J. 1994. *Student-Centered Classroom Assessment*. New York, NY: Mcmillan College Publishing Company.
- Undang-Undang RI No. 20, 2003 tentang Sistem Pendidikan Nasional*. Jakarta: Direktorat Jenderal Pendidikan Dasar dan Menengah, Bagpro Penilaian Hasil Belajar Tahap Akhir Nasional.
- Undang-Undang RI No. 14, 2005 tentang Guru dan Dosen*. Jakarta: Departemen Pendidikan Nasional Republik Indonesia.